A FUZZY EXTENSION OF MAGERIT METHODOLOGY FOR RISK ANALYSIS IN INFORMATION SYSTEMS

Eloy Vicente, Antonio Jiménez and Alfonso Mateos
Departamento de Inteligencia Artificial, Facultad de Informática, Universidad Politécnica de Madrid
Campus de Montegancedo S/N, Boadilla del Monte 28660, Madrid, SPAIN

ABSTRACT

We propose a fuzzy approach to deal with risk analysis for information systems. We extend MAGERIT methodology that valuates the asset dependencies to a fuzzy framework adding fuzzy linguistic terms to valuate the different elements (terminal asset values, asset dependencies as well as the probability of threats and the resulting asset degradation) in risk analysis. Computations are based on the trapezoidal fuzzy numbers associated with these linguistic terms and, finally, the results of these operations are translated into a linguistic term by means of a similarity function.

1. INTRODUCTION

Information Systems (IS) are composed of a set of data management elements designed to provide services and benefits in areas as far a part as public administration, industrial control, the banking or geographical and weather information.

Technological developments and the universal internet access has led to an increase in system vulnerabilities, since organizations have connected ISs to corporate and even public networks that could be accessed by non-authorized personnel unless appropriate action is not taken. Besides, people within the organization have to be trained in and aware IS support technology as technology misuse can cause disastrous failures.

On top of these vulnerabilities caused by recent technological developments, there are other traditional issues, such as integrity facilities or the safeguarding of not necessarily digital documents, on which the new technologies have also had an impact. Therefore, ISs have to be analyzed with a view to risk minimization by means of well-planned actions to protect information, processes and services from possible threats. Threats range from act of terrorism, industrial espionage, etc., or even a simple unintentional human error by an operator.

Standards promoted by the *International Organization for Standardization* (2011) on IS security suggest three-stage risk analysis and management methodologies. The *planning stage* establishes the necessary points for starting up the project, defines objectives, and identifies participants and competencies. The *analysis stage* identifies the IS assets, as well as their relations (dependencies), the threats to which they are exposed and their frequency and asset degradation levels. Finally, the *risk management stage* determines the safeguards and strategies that reduce impact and risk.

In this paper, we focus on the second stage, *analysis*. Assets are the IS or related resources, necessary for an organization's correct operation and for achieving the goals set by its manager. Assets can be data, applications, software, facilities, hardware, services...

The asset dependencies are usually represented in terms of percentages, signaling how likely the failure of an asset is to affect another. Often only a few elements (*terminal assets*), usually data or services, account for the total value of an organization's assets. The value of these assets is transferred to other assets through the established dependency relations. Thus, non-terminal assets have no intrinsic values; they accumulate their value from terminal assets.

However, the methodologies based on international standards, such us (López-Crespo et al, 2006),

MEHARI (2010), CRAMM (2003), OCTAVE (Alberts and Dorofee, 2002) and OCTAVES (Alberts and Dorofee, 2005) or NIST 800-30 (Stoneburner and Gougen, 2002), obviate the difficulty of correctly assigning asset dependencies, as well as terminal asset values or the impact on the entire system caused by the materialization of a threat to an asset. Moreover, these methodologies do not consider uncertainty concerning these assessments.

In this paper we propose a fuzzy risk analysis in IS as a solution to these deficiencies. We use the arithmetic proposed by (Xu et al, 2010), extending methodologies that valuate the asset dependencies proposed by international standards to a fuzzy framework adding fuzzy linguistic terms to valuate the different elements (terminal asset values, asset dependencies...) in risk analysis. Section 2 reviews some operations on trapezoidal fuzzy numbers and introduces a fuzzy evaluation of asset dependencies. Section 3 provides a fuzzy valuation of assets on the basis of five components is provided. Threats and asset risk impact indicators are described in Section 4. In Section 5, we introduce the similarity function used to associate a linguistic term from a set with a trapezoidal fuzzy numbers. Finally, some conclusions and future research are discussed in Section 6.

2. FUZZY VALUATION OF DEPENDENCES

Let us consider the set of trapezoidal fuzzy numbers with support in [0,1], $TF[0,1]$, i.e. $\tilde{A} = (a, b, c, d; w_{\tilde{A}})$, with $0 \le a \le \underline{b} \le c \le d \le 1$; $0 \le w_{\tilde{A}} \le 1$. We use the following arithmetic proposed in (Xu et al, 2010) in $TF[0,1]$: If $\tilde{A}_1 = (a_1, b_1, c_1, d_1; w_{\tilde{A}_1})$ and $\tilde{A}_2 = (a_2, b_2, c_2, d_2; w_{\tilde{A}_2})$, then

$$\tilde{A}_1 \oplus \tilde{A}_2 = (a_1 + a_2 - a_1a_2, b_1 + b_2 - b_1b_2, c_1 + c_2 - c_1c_2, d_1 + d_2 - d_1d_2; min\{w_{\tilde{A}_1}, w_{\tilde{A}_2}\}),$$

$$\tilde{A}_1 \otimes \tilde{A}_2 = (a_1 \times a_2, b_1 \times b_2, c_1 \times c_2, d_1 \times d_2; min\{w_{\tilde{A}_1}, w_{\tilde{A}_2}\}).$$

Let us demonstrate that both operations ($\oplus$ and $\otimes$) are well defined. Operations $\oplus$ and $\otimes$ are two internal composition laws in $TF[0,1]$ that verify the following properties: both are commutative, have a neutral element and are associative.

Note that the set $TF[0,1;1] = \{(a, b, c, d; 1) \in TF[0,1]\}$ is a subset of $TF[0,1]$ closed to operations $\oplus$ and $\otimes$. This means that confined to $TF[0,1;1]$ such operations remain internal composition laws. From now on, we will consider the framework defined by $TF[0,1;1]$ and, for convenience, use $(a, b, c, d) \equiv (a, b, c, d; 1)$ notation.

As mentioned above, the assets in IS are connected by dependency relationships, and a failure of one asset may affect other assets. The structure resulting from these dependency relationships is as shown in Fig. 1, where terminal assets (data and products or services) account for total system assets.

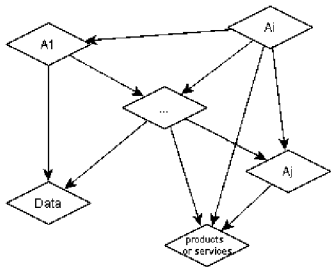

Figure 1. Asset dependencies graph.

Asset A_j *depends* on the asset A_i (or A_i *influences* A_j), denoted by (A_i, A_j) (graphically $A_i \rightarrow A_j$), if a failure in asset A_i causes a failure in the asset A_j with any given probability. This probability is usually referred to as the *degree of dependency* of A_j with respect to A_i or the *influence* of A_i over A_j, denoted by $d_d(A_i, A_j)$.

Proposed IS risk analysis methodologies (MAGERIT, MEHARI, OCTAVE...) assign just a percentage to indicate the degree of dependency between two assets, and sometimes even propose the use of a Boolean value indicating whether or not this dependency exists regardless of the degree of dependency. We propose

the use of trapezoidal fuzzy numbers to represent these dependencies. Consequently, $\tilde{d}_d(A_i, A_j) \in TF[0,1;1]$ and the experts can build a linguistic term set to intuitively define the dependency between two assets under uncertainty.

The dependency between assets in the dependency structure need not be direct but can be transitive. Namely, if (A_i, A_j) and (A_j, A_k), then (A_i, A_k). Our aim then is to compute the indirect asset dependencies since assets values are accumulated from terminal assets through these dependencies.

To avoid ambiguity (see Fig. 3), we will write D_D to refer to total dependency between two assets separated by other intermediate assets, and d_d when they are directly connected The degree of dependency of asset A_k with respect to A_i, $\tilde{D}_D(A_i, A_k)$, is computed as follows. We denote by **P**=$\{P_1,\ldots,P_s\}$ the set of paths in the analysis of the influence of A_i over A_k. Then,

A) If all assets (excluding A_i and A_k) in the paths in **P** are influenced by only one asset, then

$$\tilde{D}_D(A_i, A_k) = \oplus_{j=1}^{s} \tilde{D}_D(A_i, A_k \mid P_j), \quad (1)$$

where $\tilde{D}_D(A_i, A_k \mid P_j) = \tilde{d}_d(A_i, A_{j1}) \oplus \tilde{d}_d(A_{j1}, A_{j2}) \oplus \ldots \oplus \tilde{d}_d(A_{jn}, A_k)$ and P_j: $(A_i \rightarrow A_{j1} \rightarrow \ldots \rightarrow A_{jn} \rightarrow A_k)$.

B) Otherwise, we assume that the first *r* paths in **P** are formed by assets (excluding A_i and A_k) influenced by only one asset, and the remaining *s-r* paths include at least one asset influenced by two or more assets. Then, for the *r* first paths, we proceed as in *A*), and we denote by **S** the set including the *s-r* remaining paths. We proceed with **S** as follows:

i. Compute the set of non-terminal assets in **S** influenced by two or more assets, denoted by *I*, and the subset of *I* including assets uninfluenced by any other asset in *I*, denoted by *NI*.

ii. We consider an asset A_r in *NI*. Then, we simplify the paths in **S** that include asset A_r making $A_i \rightarrow A_r \rightarrow \ldots \rightarrow A_k$, with $\tilde{d}_d(A_i, A_r) = \tilde{D}_D(A_i, A_r)$ (computed as in *A*)).

iii. Remove repeated paths from **S** and keep only one instance

iv. Build *I* and *NI* again from **S**.

v. If *NI* is not empty, go to ii). Otherwise, the algorithm finishes.

Let us denote the resulting set of paths by **S**= $\{P'_1,\ldots,P'_m\}$, with $m \leq s\text{-}r$. Then, the degree of dependency of A_k regarding A_I is $\tilde{D}_D(A_i, A_k) = \oplus_{j=1}^{r} \tilde{D}_D(A_i, A_k \mid P_j) \oplus_{l=1}^{m} \tilde{D}_D(A_i, A_k \mid P'_l)$.

Fig. 2 shows an example of dependency structure in IS. The degree of dependence of A_6 with respect to A_1 is computed as follows. First, **P** = $\{P_1$: $(A_1 \rightarrow A_2 \rightarrow A_6)$, P_2: $(A_1 \rightarrow A_2 \rightarrow A_3 \rightarrow A_6)$, P_3: $(A_1 \rightarrow A_2 \rightarrow A_3 \rightarrow A_4 \rightarrow A_6)$, P_4: $(A_1 \rightarrow A_3 \rightarrow A_6)$, P_5: $(A_1 \rightarrow A_3 \rightarrow A_4 \rightarrow A_6)$, P_6: $(A_1 \rightarrow A_4 \rightarrow A_6)$, P_7: $(A_1 \rightarrow A_5 \rightarrow A_6)\}$.

Asset A_3 is influenced by A_1 and A_2, and A_4 is influenced by A_1 and A_3. Therefore, we apply *B*) with $r = 2$ and **S** = $\{P_2, P_3, P_4, P_5, P_6\}$ and proceed as follows:

i. $I = \{A_3, A_4\}$ and $NI = \{A_3\}$.

ii. We select A_3, then we simplify the paths P_2, P_3, P_4 and P_5 to P'_2: $(A_1 \rightarrow A_3 \rightarrow A_6)$, P'_3: $(A_1 \rightarrow A_3 \rightarrow A_4 \rightarrow A_6)$, P'_4: $(A_1 \rightarrow A_3 \rightarrow A_6)$ and P'_5: $(A_1 \rightarrow A_3 \rightarrow A_4 \rightarrow A_6)$, respectively, with

$$\tilde{d}_d(A_1, A_3) = \tilde{D}_D(A_1, A_3) = (\tilde{d}_d(A_1, A_2) \otimes \tilde{d}_d(A_2, A_3)) \oplus \tilde{d}_d(A_1, A_3).$$

iii. **S** = $\{P'_2, P'_3, P_6\}$, since $P'_2 = P'_4$ and $P'_3 = P'_5$.

iv. $I = \{A_4\}$ and $NI = \{A_4\}$.

v. Go to (ii).

ii. We select A_4, then we simplify the paths P'_3 and P_6 to P''_3: $(A_1 \rightarrow A_4 \rightarrow A_6)$ and P'_6:$(A_1 \rightarrow A_4 \rightarrow A_6)$, respectively, with $\tilde{d}_d(A_1, A_4) = \tilde{D}_D(A_1, A_4) = (\tilde{d}_d(A_1, A_3) \otimes \tilde{d}_d(A_3, A_4)) \oplus \tilde{d}_d(A_1, A_4)$.

iii. **S** = $\{P'_2, P''_3\}$, since $P''_3 = P'_6$.

iv. $I = \varnothing$ and $NI = \varnothing$.

v. The algorithm finishes since $NI = \varnothing$.

Finally, **S** = $\{P'_2, P''_3\}$ and the degree of dependence of A_6 with respect to A_1 is

$$\begin{aligned}\tilde{D}_D(A_1,A_6) &= \tilde{D}_D(A_1,A_6 \mid P_1) \oplus \tilde{D}_D(A_1,A_6 \mid P_7) \oplus \tilde{D}_D(A_1,A_6 \mid P_2') \oplus \tilde{D}_D(A_1,A_6 \mid P_3'') \\ &= (\tilde{d}_d(A_1,A_2) \otimes \tilde{d}_d(A_2,A_6)) + (\tilde{d}_d(A_1,A_5) \otimes \tilde{d}_d(A_5,A_6)) + \\ &+ (\tilde{d}_d(A_1,A_3) \otimes \tilde{d}_d(A_3,A_6)) + (\tilde{d}_d(A_1,A_4) \otimes \tilde{d}_d(A_4,A_6))\end{aligned}$$

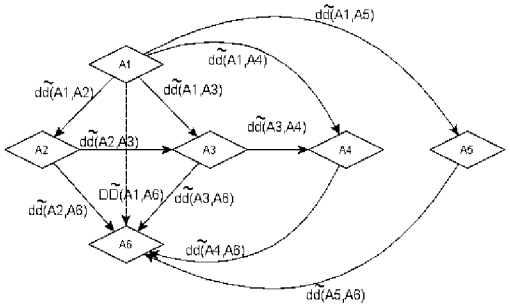

Figure 2. Asset dependencies structure in ISs.

Note that transactions between trapezoidal fuzzy numbers representing linguistic terms from a set in [0,1] will remain in $TF[0,1;1]$, and the results of these operations can be translated into one of the linguistic terms of the set by means of a *similarity function*. Furthermore, the operation $\oplus$ is consistent with the methodologies established for risk analysis and management in IS, allowing performances in probabilistic terms.

Let us consider these issues in more detail. The MAGERIT (López-Crespo et al, 2006) methodology uses real numbers to determine both the dependency and the asset values, as mentioned before, and proposes the operation $a \overline{\oplus} b = a + b - ab$.

Operation $\overline{\oplus}$ is a special case of the operation proposed in (Xu et al, 2010) and used in this paper, $\oplus$, since a real number $a \in [0,1]$ can be written as the trapezoidal fuzzy number $\tilde{a} = (a, a, a, a)$, and therefore $\tilde{a} \oplus \tilde{b} = (a,a,a,a) \oplus (b,b,b,b) = (a+b-ab, a+b-ab, a+b-ab, a+b-ab) = a \tilde{\overline{\oplus}} b \approx a \overline{\oplus} b$.

On the other hand, operation $\otimes$ extends naturally to the product of real numbers. Therefore, by defining operations $\oplus$ and $\otimes$, we have successfully extended the basic operations using IS risk analysis and management methodologies to the context of fuzzy numbers.

Probability theory is the source of the second reason for using the operation $\oplus$. To illustrate this point, we shall use the examples in Fig. 3. In the first example, a failure in asset A_1 has the given probabilities of causing a failure in asset A_4 by means of A_2 or A_3. As both paths are not necessarily mutually exclusive, the computation of probabilities is $p = p_1 p_3 + p_2 p_4 - p_1 p_3 p_2 p_4 = p_1 p_3 \overline{\oplus} p_2 p_4 \approx [\tilde{p}_1 \otimes \tilde{p}_3] \oplus [\tilde{p}_2 \otimes \tilde{p}_4]$, where $\tilde{a} = (a, a, a, a) \approx a$ is the fuzzy representation of the real number a.

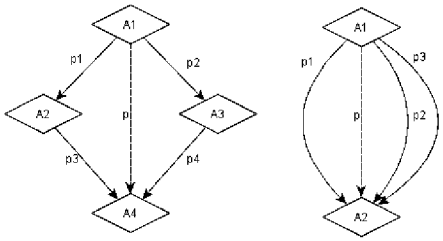

Figure 3. Asset dependencies structure with two and three paths.

Suppose that we have three instead of two alternative paths, see the second example in Fig. 3, and we assume that the total probabilities for these paths are p_1, p_2 and p_3 (for convenience, we disregard the intermediate assets). Thus, according to probability theory, $p = p_1 + p_2 + p_3 - p_1p_2 - p_1p_3 - p_1p_2p_3 = p_1 \overline{\oplus} p_2 \overline{\oplus} p_3$, and, by induction on the number of paths between A_1 and A_2, we can infer that for a set of real numbers

$$p_1, p_2, ..., p_n \in [0,1],\ p_1 \bar{\oplus} p_2 \bar{\oplus} ... \bar{\oplus} p_n = \sum_{i=1}^{n} p_i - \sum_{i<j} p_i p_j + \sum_{i<j<k} p_i p_j p_k + ... + (-1)^{n+1} \prod_{i=1}^{n} p_i.$$

This expression can obviously be extended to fuzzy numbers in $TF[0,1;1]$, componentwise, with the sum and product operations ($\oplus$ and $\otimes$). Given these operations, the methodology for deriving the dependency from each terminal asset is: First, we establish a set of fuzzy linguistic terms. Then, the experts identify the degree of dependency for each pair of consecutive assets in the general dependency structure using linguistic terms in the above set. Finally, the degree of dependency on assets with respect to terminal assets is computed using Eqs. (1) or (2).

As an example, let us consider the dependency structure in Fig. 4, including the degree of dependency for each pair of assets on the basis of the linguistic term shown in Table 1. Table 2 shows the degree of dependency of nonterminal assets with respect to A_6.

Table 1. Linguistic term set

Term	Fuzzy Number
Very low (VL)	(0, 0, 0, 0.05)
Low (L)	(0, 0.075, 0.125, 0.275)
Medium-Low (M-L)	(0.125, 0.275, 0.325, 0.475)
Medium (M)	(0.325, 0.475, 0.525, 0.675)
Medium-High (M-H)	(0.525, 0.675, 0.725, 0.875)
High (H)	(0.725, 0.875, 0.925, 1)
Very High (VH)	(0.925, 1, 1, 1)

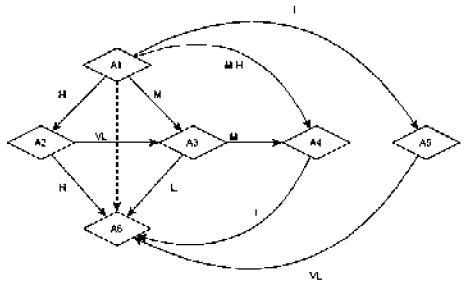

Figure 4. Asset dependencies.

Table 2. Accumulated degree of dependency for non-terminal assets

A_i	$\tilde{D}_D(A_i, A_6)$
A_1	(0.679, 0.891, 0.949, 1)
A_2	(0.725, 0.875, 0.925, 1)
A_3	(0, 0.107, 0.182, 0.409)
A_4	(0, 0.075, 0.125, 0.275)
A_5	(0, 0, 0, 0.05)

3. FUZZY SETS VALUATIONS

MAGERIT defines the *value of an asset* as the losses that would be sustained if the respective asset is no longer available. These can be losses of money, user confidence, the organizational prestige... Assets are usually evaluated taking into account the following five components (López-Crespo et al, 2006):

- *Confidentiality*. How much damage would it cause if the asset is disclosed to someone it should not be?
- *Integrity*. How much damage would it cause if the asset is damaged or corrupt? Data can be manipulated, be wholly or partially false, or even missing.
- *Authenticity*. How much damage would it cause if we do not exactly know who has done what? This is a typical services (user authentication) and data (authenticity of the person accessing data inspection.

- *Traceability*. How much damage would it cause if it is not known for whom the service is being provided?, i.e. who does what and when? How much damage would it cause if it is not known who accessed what data and what they did with them?
- *Availability*. How much damage would it cause if the asset is not available or cannot be used?

Confidentiality, integrity, and availability are typical data inspections. Only the terminal assets have an associated value for the above components. The other assets accumulate value from terminal assets on the basis of dependency relationships. We again use the set of linguistic terms that represent trapezoidal fuzzy numbers to represent uncertainty when valuating the terminal assets.

Let us denote assets by $\tilde{\mathbf{v}}_j = (\tilde{v}_{j_{(1)}}, \tilde{v}_{j_{(2)}}, \tilde{v}_{j_{(3)}}, \tilde{v}_{j_{(4)}}, \tilde{v}_{j_{(5)}})$, where $\tilde{v}_{j_{(l)}}$ is a linguistic term assigned by an expert for the *l*th value component in asset A_j. If we denote by TAS the terminal asset set, then the value of asset A_j with respect to terminal assets is $\tilde{v}_{j_{(l)}} = \otimes_{A_k \in TAS}(\tilde{D}_D(A_j, A_k) \otimes \tilde{v}_{k_{(l)}})$.

If the values provided by expert for the five components in the terminal asset A_6 are confidentiality (H), integrity (H), authenticity (M), traceability (L) and availability (H), then the accumulated values for the non-terminal asset A_1 are shown in Table 3.

Table 3. Accumulated values for A_1

Component	$\tilde{v}_{1_{(l)}}$
Confidentiality	(0.492, 0.779, 0.877, 1)
Integrity	(0.492, 0.779, 0.877, 1)
Authenticity	(0.220, 0.423, 0.498, 0.675)
Traceability	(0, 0.066, 0.118, 0.275)
Availability	(0.492, 0.779, 0.877, 1)

4. THREATS

Next, we assess threats and estimate indicators of the impact on and risk to assets. A *threat* is an event that can trigger an incident in our organization, causing damage or intangible material loss to the assets, and an *attack* is any deliberate action aimed at violating the IS security mechanisms. MAGERIT suggests two threat assessment measures: *degradation*, the damage that the threat can cause to the asset, and *frequency*, how often the threat materializes.

We will again use fuzzy linguistic terms rather than percentages and probabilities to represent degradation and frequency. A threat is a vector $\tilde{\tilde{u}} = (\tilde{D}, \tilde{f})$ whose components are degradation and frequency. Note that the degradation has to be established for each the five asset components described in the previous section, $\tilde{\tilde{D}} = (\tilde{d}_1, \tilde{d}_2, \tilde{d}_3, \tilde{d}_4, \tilde{d}_5)$ i.e., the threat causes a degradation $\tilde{d}_l$ in the *l*th component of the asset.

Let us consider a threat on the asset A_j. When the threat is realized, each component is affected by the expression $\tilde{I}_{j_{(l)}} = \tilde{d}_l \otimes \tilde{v}_{j_{(l)}}$, where $\tilde{I}_{j_{(l)}}$ is the impact on the *l*th component of the attacked asset (A_j).

We compute the *risk* to the attacked asset by the expression $\tilde{R}_{j_{(l)}} = \tilde{I}_{j_{(l)}} \otimes \tilde{f}$. After computing the impact caused by a materialized threat on an asset, we can compute the impact transmitted from the attacked asset to its dependent assets. If A_j is the asset on which the threat has materialized and the degree of dependency of A_j with respect to A_k is $\tilde{D}_D(A_k, A_j)$, then the attack on asset A_j has an impact on A_k of $\tilde{I}_{k_{(l)}} = \tilde{D}_D(A_k, A_j) \otimes \tilde{d}_l \otimes \tilde{v}_{j_{(l)}}$. Thus, the risk to asset A_k is $\tilde{R}_{k_{(l)}} = \tilde{I}_{k_{(l)}} \otimes \tilde{f} = \tilde{D}_D(A_k, A_j) \otimes \tilde{d}_l \otimes \tilde{v}_{j_{(l)}} \otimes \tilde{f}$.

We consider a threat on asset A_3, see Fig. 4, with a degradation $\tilde{D}$ = (H, L, M, VL, M) and frequency M, respectively. Then, the impact on and the risk to asset A_3 are shown in Table 4.

The impacts on and risk to the assets depending on A_3, i.e. A_1 and A_2, are now computed, see Table 4. Note that the low impacts on and risks to A_2 are due to the dependency between A_2 and A_3 is very low (VL), see Fig. 4, with an associated trapezoidal fuzzy number (0, 0, 0, 0.05).

Table 4. Impact and risk to A_3

Component	Impact on A_3	Risk to A_1	Risk to A_2	Risk to A_3
Confidentiality	(0, 0.081, 0.155, 0.409)	(0.037, 0.143, 0.223, 0.466)	(0, 0, 0, 0.033)	(0, 0.038, 0.081, 0.276)
Integrity	(0, 0.007, 0.021, 0.112)	(0, 0.013, 0.03, 0.128)	(0, 0, 0, 0.009)	(0, 0.003, 0.011, 0.075)
Authenticity	(0, 0.024, 0.050, 0.186)	(0.007, 0.045, 0.072, 0.212)	(0, 0, 0, 0.015)	(0, 0.011, 0.026, 0.125)
Traceability	(0, 0, 0, 0.005)	(0, 0, 0, 0.006)	(0, 0, 0, 0.0004)	(0, 0, 0, 0.038)
Availability	(0, 0.044, 0.088, 0.27)	(0, 0.013, 0.03, 0.128)	(0, 0, 0, 0.022)	(0, 0.021, 0.046, 0.186)

5. SIMILARITY FUNCTION

A *similarity function* is required to associate the resulting trapezoidal fuzzy number with an element in the linguistic term set. This function can also be used at any step of the methodology to derive the linguistic terms associated with the respective trapezoidal fuzzy numbers output to represent dependencies, accumulated values...

Several authors have proposed different similarity functions, which are based on the centroid of a fuzzy number and the distance between the components of the fuzzy numbers in $TF[0,1]$, see (Hsieh and Chen, 1999; Chen and Chen 2003, 2007). A more recent similarity function was proposed in (Xu et al, 2010) and compared with the proposal reported in (Chen and Chen, 2007).

However, the above similarity functions are unsuitable for use in $TF[0,1;1]$. We use the function proposed in (Vicente et al, 2012), which considers another parameter consisting of the ratio between the common area and the joint area under the membership functions of trapezoidal fuzzy numbers. Moreover, we use the distance l_∞ between centroids since the use of distances with nonrectangular spheres is inconsistent with the intuitive perception of similarity in $TF[0,1;1]$. Given $\tilde{A}$ and $\tilde{B} \in TF[0,1;1]$, the similarity function can be defined as

$$S(\tilde{A},\tilde{B}) = 1 - w_1\left(1 - \frac{\int_0^1 \mu_{\tilde{A}\cap\tilde{B}}(x)dx}{\int_0^1 \mu_{\tilde{A}\cup\tilde{B}}(x)dx}\right) - w_2\frac{\sum|a_i - b_i|}{4} - w_3 l_\infty[(X_{\tilde{A}}, Y_{\tilde{A}}),(X_{\tilde{B}}, Y_{\tilde{B}})],$$

where $w_1 + w_2 + w_3 = 1$, $(X_{\tilde{A}}, Y_{\tilde{A}})$ and $(X_{\tilde{B}}, Y_{\tilde{B}})$ are the centroids of $\tilde{A}$ and $\tilde{B}$, respectively, i.e.

$$X_{\tilde{A}} = Y_{\tilde{A}}(a_3 + a_2) + (1 - Y_{\tilde{A}})(a_4 + a_1) \quad \text{and} \quad Y_{\tilde{A}} = \begin{cases} \left(\frac{a_3 - a_2}{a_4 - a_1} + 2\right)/6, & \text{if } a_4 - a_1 \neq 0 \\ 1/2, & \text{if } a_4 - a_1 = 0 \end{cases},$$

$\mu_{\tilde{A}\cap\tilde{B}} = \min_{0\le x\le 1}\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}$, $\mu_{\tilde{A}\cup\tilde{B}} = \max_{0\le x\le 1}\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}$, with μ_χ the membership function of χ, and $l_\infty((x_1, y_1), (x_2, y_2)) = \max\{|x_1 - x_2|, |y_1 - y_2|\}$.

Note that w_1, w_2 and w_3 represent the relative importance of the three elements considered in the similarity function. Analysts will assign the values that best fits their own model.

Looking at the same example than in previous sections, applying the similarity function with equal weights for all three components, then the corresponding linguistic terms for the risk to assets A_1, A_2 and A_3 are shown in Table 6. Fig. 5 shows the risk components to A_1.

6. CONCLUSIONS

We have developed a fuzzy risk analysis model for information systems that conforms to international standards, particularly the MAGERIT methodology. The model is an improvement on this and other existing methodologies since it includes uncertainty about the assessments by means of linguistic terms, which have associated trapezoidal fuzzy numbers. The proposed methodology makes computations on the basis of trapezoidal fuzzy numbers to accumulate dependencies between assets and asset valuations and to determine impacts and risk from the threat degradation and frequency, respectively. Moreover, similarity functions can be used at any step in the methodology to derive a linguistic term for the trapezoidal fuzzy number output.

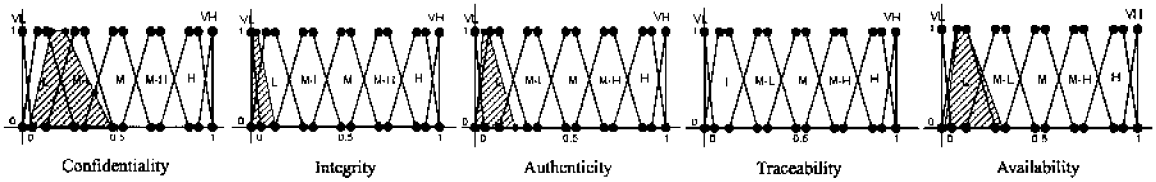

Figure 5. Risk components

Table 5. Linguistic terms for the risk to A_1, A_2 and A_3

Component	A_1	A_2	A_3
Confidentiality	M-L	VL	L
Integrity	VL	VL	VL
Authenticity	L	VL	VL
Traceability	VL	VL	VL
Availability	L	VL	VL

ACKNOWLEDGEMENT

The paper was supported by Madrid Regional Government project S-2009/ESP-1685 and the Spanish Ministry of Science and Innovation project MTM2011-28983-C03-03.

REFERENCES

Alberts, C. and Dorofee, A., 2002. *Managing Information Security Risks: The OCTAVE Approach*. Addison- Wesley, New York, USA.

Alberts, C. and Dorofee, A., 2005. *OCTAVE-s Method Implementation Guide Version 2.0*. Canergie Mellon University, Pittsburgh, USA.

Chen, S.-J. and Chen, S.-M., 2003. Fuzzy Risk Analysis Based on Similarity Measures of Generalized Fuzzy Numbers. *In IEEE Transactions on Fuzzy Systems*, Vol. 11, pp. 45-56.

Chen, S.-J. and Chen, S.-M., 2007. Fuzzy Risk Analysis Based on the Ranking of Generalized Trapezoidal Fuzzy Numbers. *In Applied Intelligence*, Vol. 26, pp. 1-11.

CCTA Risk Analysis and Management Method (CRAMM), Version 5.0. (2003). Central Computing and Telecommunications Agency (CCTA), London, UK.

Hsieh, C.H. and Chen, S.H., 1999. Similarity of Generalized Fuzzy Numbers with Graded Mean Integration Representation. *Proceedings of the 8th International Fuzzy Systems Association World Congress*, pp. 551-555.

ISO/IEC 27005:2011, Information technology – Security techniques - Information security risk management. (2011). International Organization for Standarization, Geneva, Switzerland.

López-Crespo F. et al, 2006. *Methodology for Information Systems Risk Analysis and Management (MAGERIT version 2). Book III-The Techniques*. Ministerio de Administraciones Pública, Madrid, España.

Mehari 2010 - Risk Analysis and Treatment Guide (2010). Club de la Sécurité de l'Information Francais (CSIF), Paris, France.

Stoneburner, G. and Gougen, A., 2002. *NIST 800- 30 Risk Management. Guide for Information Technology Systems*. National Institute of Standard and Technology, Gaithersburg, USA.

Vicente E. et al., 2012. A New Similarity Measure of Trapezoidal Fuzzy Numbers. *In Expert Systems with Applications*, under review.

Xu Z. et al., 2010. A Method for Fuzzy Risk Analysis based on the New Similarity of Trapezoidal Fuzzy Numbers. *In Expert Systems with Applications*, Vol. 37, pp. 1920-1927.